Mapping of the Trichohyalin Gene: Co-Localization with the Profilaggrin, Involucrin, and Loricrin Genes

Michael J. Fietz, George E. Rogers, Helen J. Eyre, Elizabeth Baker, David F. Callen, and Grant R. Sutherland

Department of Biochemistry (MJF, GER), University of Adelaide; and Department of Cytogenetics and Molecular Genetics (HJE, EB, DFC, GRS), Adelaide Children's Hospital, North Adelaide, South Australia, Australia

The chromosomal location of the gene encoding the human hair follicle protein trichohyalin has been determined by *in situ* hybridization. The human gene has been localized to the region 1q21.1–1q23 (probably 1q21.3) using a sheep trichohyalin cDNA probe. The genes encoding three other epithelial proteins, namely, profilaggrin, involucrin, and loricrin, are also located in the same region of chromosome 1, which, together with their similar gene and protein structures, suggests that the four proteins form a novel superfamily of epithelial structural proteins. *J Invest Dermatol 99:542–544, 1992*

The cytoskeleton of most types of eukaryote cells includes the intermediate filaments (IF) that are formed from a large set of sequence-related proteins. On the basis of their sequences the IF proteins have been subdivided into five types, of which types I and II are keratins and are specifically expressed within the epithelial tissues [1]. Once formed, the IF are associated with numerous other proteins, termed intermediate filament–associated proteins (IFAP), that organize the arrangement of the IF within the cell [1,2].

Analysis of gene mapping and chromosomal localization data has shown that the genes encoding IF and IFAP protein families are clustered within the genome. Thus the genes for the localized type I IF proteins (apart from one) have been mapped to human chromosome 17 [3–6], whereas the type II genes are located on human chromosome 12 [4,5,7,8] with all of the finely mapped type II genes clustered in the region 12q11–12q14. Of the genes encoding hair IFAP, it is presently known that at least two ultrahigh-sulphur IFAP are located on human chromosome 11 [9], whereas two sheep high-sulphur IFAP also are known to be clustered within the sheep genome [10].

Another group of clustered genes consists of those encoding the epidermal structural proteins profilaggrin, involucrin, and loricrin. These proteins do not have identical functions but they or their processed form, as is the case with profilaggrin (filaggrin), are all incorporated into the cross-linked epidermal cell envelope by the enzyme transglutaminase [11–14] with filaggrin also playing a major role in the aggregation of the epidermal IF network [15,16]. Sequence analysis has shown that each of these proteins contains a series of tandem peptide repeats [14,17–19]. The genes encoding these three proteins, which are all located at 1q21 in the human genome [20–22], also share a common feature in that they have only a single intron that is present in the 5′ non-coding region [19,22]. Another epidermal protein with similar characteristics to those described above for profilaggrin, involucrin, and loricrin is the hair follicle protein trichohyalin. Trichohyalin is incorporated into the IF-like structures in the mature cells of the hair follicle inner root sheath [23]. Trichohyalin, like filaggrin, involucrin, and loricrin, becomes highly cross-linked by the enzyme transglutaminase. The trichohyalin molecule has a mass of approximately 200 kd and consists predominantly of tandem repeats of an approximately 23–amino-acid sequence [24]. Although the gene for trichohyalin has an intron present in the 5′ non-coding region, as is the case for the profilaggrin, involucrin, and loricrin genes, it also contains an intron within the 5′ end of the coding region [24].

Given the epidermal origin of trichohyalin and the structure and properties of this gene, relative to profilaggrin, involucrin, and loricrin, the location of the trichohyalin gene within the human genome has been determined.

MATERIALS AND METHODS

Genomic Southern analysis was performed as previously described [25]. The probe was radiolabeled by the method of Feinberg and Vogelstein [26] using the Hexaprime DNA Labelling Kit from Bresatec Limited (Adelaide, Australia).

The *in situ* hybridization procedure was previously described [27]. The probe was tritium-labeled to a specific activity of approximately 7.2×10^7 cpm/μg and hybridized to metaphases from two normal men at a concentration of 0.2–0.4 μg/ml and exposed for 17–27 d. All silver grains touching chromosomes were counted to determine the pattern of hybridization of the probe.

RESULTS

In earlier studies, a sheep cDNA clone, corresponding to approximately 40% of the full-length trichohyalin mRNA, was purified and sequenced [25]. To examine the homology with the human trichohyalin gene, a 1.9-kb *Eco*R I fragment from the cDNA clone, a portion of which encodes a segment of the repetitive region, was used to probe a genomic blot containing both sheep and human genomic DNA that had been digested with *Eco*R I (Fig 1). The probe hybridized to a single fragment in the human DNA with an intensity of approximately 25% of that obtained with the corresponding sheep fragment, indicating that the sheep and human genes show considerable homology. Therefore a clone containing the 1.9-kb sheep cDNA fragment was used in an attempt to localize the human trichohyalin gene.

The *in situ* hybridization results (Fig 2) showed that, out of 115 silver grains on 25 metaphases, 17 (14.8% of all silver grains) were on the long arm of chromosome 1. The distribution of the silver grains on an additional 25 metaphases in which there were silver grains on 1q are shown in Fig 3. These data localize trichohyalin to

Manuscript received April 22, 1992; accepted for publication June 30, 1992.

Reprint requests to: Prof. George Rogers, Department of Biochemistry, University of Adelaide, GPO Box 498, Adelaide, S.A., 5001, Australia.

Abbreviations:

IF: intermediate filament

IFAPs: intermediate filament–associated protein

0022-202X/92/$05.00

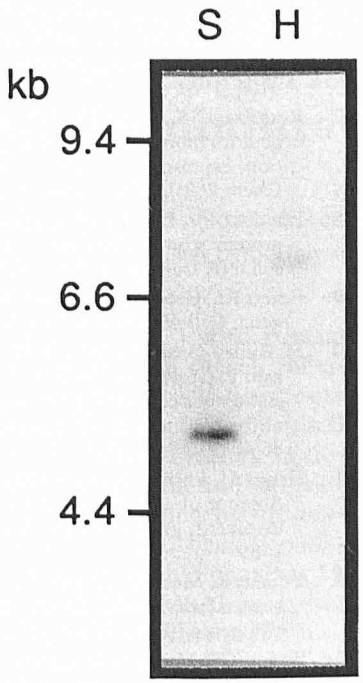

Figure 1. Southern blot analysis of sheep and human genomic DNA. Total sheep (lane S) and human (H) genomic DNA samples (4 μg/lane) were digested with *Eco*RI and separated on a 1% agarose gel. The resultant filter was probed with a 1.9-kb sheep trichohyalin cDNA fragment [25] that hybridized to only a single band in each of the tracks. Size markers are shown (in kilobase pairs).

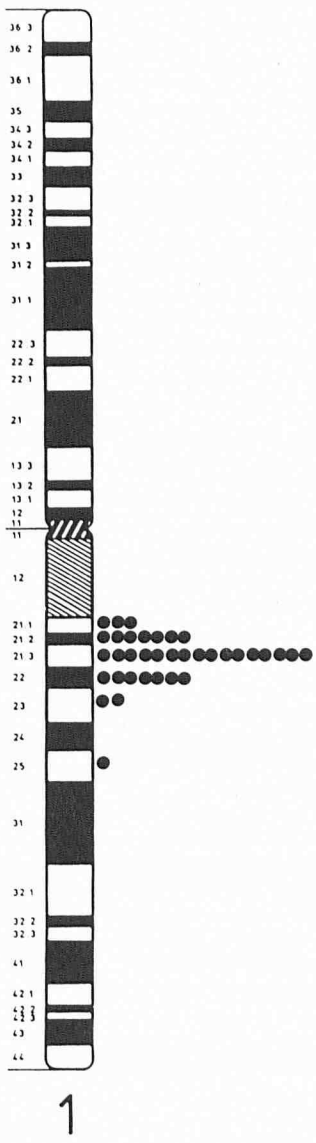

Figure 3. Distribution of silver grains over the long arm of chromosome 1 from 25 metaphases showing signal on this chromosome arm.

the region 1q21.1–1q23 but most probably to 1q21.3. A similar result was obtained from the hybridization to the metaphases from the second man (data not shown).

DISCUSSION

The localization of the trichohyalin gene at position 1q21 with the genes encoding profilaggrin, involucrin, and loricrin, as determined by chromosomal in situ hybridization using a tritium-labeled sheep cDNA probe, is of great interest given similarities that exist between the four proteins and their corresponding genes. All four proteins contain a large tandem array of peptide repeats, are transglutaminase substrates, and are involved in the formation of one or more of the highly cross-linked structures of the epidermis or hair follicle. Each of the genes is notable by the presence of an intron in the 5′ non-coding region as well as by the absence of introns in the coding region, except for trichohyalin, which has a single intron in the coding region, 138 bp from the initiating ATG. Comparison of the gene sequences from a number of mammalian species reveals that the peptide repeat sequences within each of the four genes have been poorly conserved during evolution [18,28–31], suggesting that the sequences in the individual genes have been subject to rapid alteration. Although the four genes are markedly different in sequence, it therefore appears that they may encode a superfamily of proteins originating from a single precursor. During evolution, the

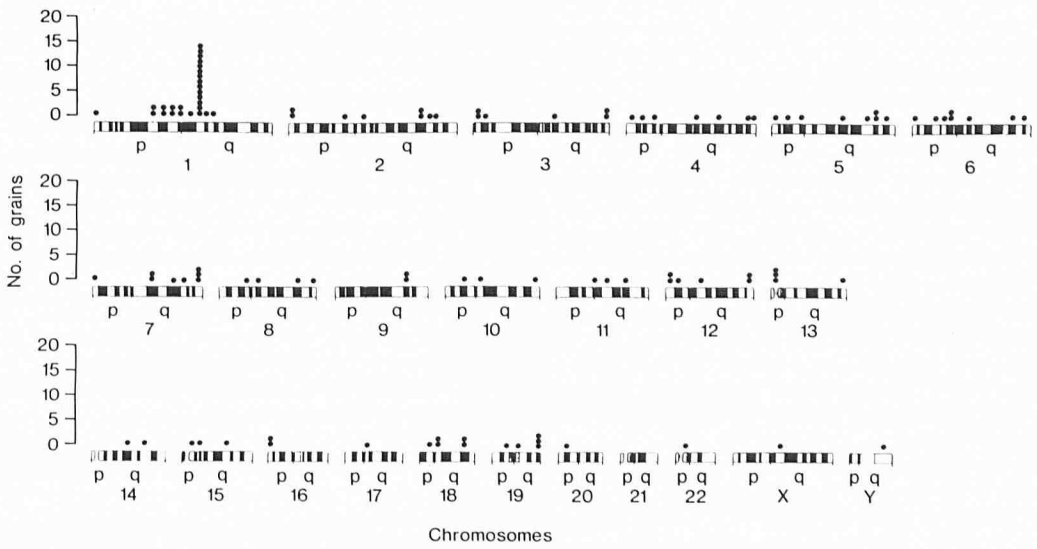

Figure 2. Distribution of silver grains over all chromosomes in 25 metaphases after in situ hybridization with the trichohyalin probe.

gene structure and protein function have remained similar although the actual nucleotide sequences have diverged significantly. On the basis of these conclusions, it is possible that the genes encoding other epidermal structural proteins will also be found to share the same gene structure and chromosomal location as the aforementioned genes. One possible example is the gene encoding the epidermal transglutaminase substrate keratolinin whose cDNA sequence has been recently determined [32].

The possibility that trichohyalin and profilaggrin belong to the same protein family may shed some light on the role of trichohyalin. It has long been questioned whether the proteins of the trichohyalin granules form the filaments of the hardened hair follicle inner root sheath or whether they act as an IFAP [25,33–38]. On the basis of filaggrin's role as the epidermal IFAP, it therefore seems more likely that trichohyalin acts as an IFAP in the inner root sheath, in accord with conclusions drawn from earlier immunoelectron microscopy data [38].

This work was supported by a grant from the Australian Research Council.

REFERENCES

1. Steinert PM, Roop DR: Molecular and cellular biology of intermediate filaments. Ann Rev Biochem 57:593–625, 1988
2. Goldman RD, Goldman AE, Green K, Jones JCR, Lieska N, Yang S-Y: Intermediate filaments—possible functions as cytoskeletal connecting links between the nucleus and the cell-surface. Ann NY Acad Sci 455:1–17, 1985
3. Bader BL, Jahn L, Franke WW: Low level expression of cytokeratins 8, 18 and 19 in vascular smooth muscle cells of human umbilical cord and in cultured cells derived therefrom, with an analysis of the chromosomal locus containing the cytokeratin 19 gene. Eur J Cell Biol 47:300–319, 1988
4. Lessin SR, Huebner K, Isobe M, Croce CM, Steinert PM: Chromosomal mapping of human keratin genes: evidence of non-linkage. J Invest Dermatol 91:572–578, 1988
5. Romano V, Bosco P, Rocchi M, Costa G, Leube RE, Franke WW, Romeo G: Chromosomal assignments of human type I and type II cytokeratin genes to different chromosomes. Cytogenet Cell Genet 48:148–151, 1988
6. Rosenberg M, RayChaudhury A, Shows TB, Le Beau MM, Fuchs E: A group of type I keratin genes on human chromosome 17: characterization and expression. Mol Cell Biol 8:722–736, 1988
7. Popescu NC, Bowden PE, DiPaolo JA: Two type II keratin genes are localized on human chromosome 12. Hum Genet 82:109–112, 1989
8. Rosenberg M, Fuchs E, Le Beau MM, Eddy RL, Shows TB: Three epidermal and one simple epithelial type II keratin genes map to human chromosome 12. Cytogenet Cell Genet 57:33–38, 1991
9. MacKinnon PJ, Powell BC, Rogers GE, Baker EG, MacKinnon RN, Hyland VJ, Callen DF, Sutherland GR: An ultrahigh-sulphur keratin gene of the human hair cuticle is located at 11q13 and cross-hybridizes with sequences at 11p15. Mammalian Genome 1:53–56, 1991
10. Powell BC, Sleigh MJ, Ward KA, Rogers GE: Mammalian keratin gene families: organisation of genes coding for the B2 high-sulphur proteins of sheep wool. Nucleic Acids Res 11:5327–5346, 1983
11. Rice RH, Green H: Presence in human epidermal cells of a soluble protein precursor of the cross-linked envelope: activation of the cross-linking by calcium ions. Cell 18:681–694, 1979
12. Simon M, Green H: Participation of membrane-associated proteins in the formation of the cross-linked envelope of the keratinocyte. Cell 36:827–834, 1984
13. Richards S, Scott IR, Harding CR, Liddell JE, Powell GM, Curtis CG: Evidence for filaggrin as a component of the cell envelope of the newborn rat. Biochem J 253:153–160, 1988
14. Mehrel T, Hohl D, Rothnagel JA, Longley MA, Bundman D, Cheng C, Lichti U, Bisher ME, Steven AC, Steinert PM, Yuspa SH, Roop DR: Identification of a major keratinocyte cell envelope protein, loricrin. Cell 61:1103–1112, 1990
15. Dale BA, Holbrook KA, Steinert PM: Assembly of *stratum corneum* basic protein and keratin filaments in macrofibrils. Nature (Lond) 276:729–731, 1978
16. Steinert PM, Cantieri JS, Teller DC, Lonsdale-Eccles JD, Dale BA: Characterization of a class of cationic proteins that specifically interact with intermediate filaments. Proc Natl Acad Sci USA 78:4097–4101, 1981
17. Rothnagel JA, Mehrel T, Idler WW, Roop DR, Steinert PM: The gene for mouse epidermal filaggrin precursor: its partial characterization, expression, and sequence of a repeating filaggrin unit. J Biol Chem 262:15643–15648, 1987
18. Haydock PV, Dale BA: Filaggrin, an intermediate filament-associated protein: structural and functional implications from the sequence of a cDNA from rat. DNA Cell Biol 9:251–261, 1990
19. Eckert RL, Green H: Structure and evolution of the human involucrin gene. Cell 46:583–589, 1986
20. McKinley-Grant LJ, Idler WW, Bernstein IA, Parry DAD, Cannizzaro L, Croce CM, Huebner K, Lessin SR, Steinert PM: Characterization of a cDNA clone encoding human filaggrin and localization of the gene to chromosome region 1q21. Proc Natl Acad Sci USA 86:4848–4852, 1989
21. Simon M, Phillips M, Green H, Stroh H, Glatt K, Bruns G, Latt SA: Absence of a single repeat from the coding region of the human involucrin gene leading to RFLP. Am J Hum Genet 45:910–916, 1989
22. Yoneda K, McBride W, Idler W, Hohl D, Steinert P: Structure of the human loricrin gene: linkage at 1q21 with profilaggrin and involucrin genes (abstr). J Invest Dermatol 96:541, 1991
23. Rogers GE, Harding HWJ, Llewellyn-Smith IJ: The origin of citrulline-containing proteins in the hair follicle and the chemical nature of trichohyalin, an intracellular precursor. Biochim Biophys Acta 495:159–175, 1977
24. Rogers GE, Fietz MJ, Fratini A: Trichohyalin and matrix proteins. Ann NY Acad Sci 642:64–81, 1991
25. Fietz MJ, Presland RB, Rogers GE: The cDNA-deduced amino acid sequence for trichohyalin, a differentiation marker in the hair follicle, contains a 23 amino acid repeat. J Cell Biol 110:427–436, 1990
26. Feinberg AD, Vogelstein B: A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132:6–13, 1983
27. Sutherland GR, Baker E, Callen DF, Campbell HD, Young IG, Sanderson CJ, Garson OM, Lopez AF, Vadas MA: Interleukin-5 is at 5q31 and is deleted in the 5q- syndrome. Blood 71:1150–1152, 1988
28. Tseng H, Green H: Remodeling of the involucrin gene during primate evolution. Cell 54:491–496, 1988
29. Djian P, Green H: Vectorial expansion of the involucrin gene and the relatedness of the hominoids. Proc Natl Acad Sci USA 86:8447–8451, 1989
30. Fietz MJ: Purification and analysis of the trichohyalin gene: an examination of the role of trichohyalin in the inner root sheath. Ph.D. thesis, University of Adelaide, Adelaide, 1991
31. Hohl D, Mehrel T, Lichti U, Turner ML, Roop DR, Steinert PM: Characterization of human loricrin: structure and function of a new class of epidermal cell envelope proteins. J Biol Chem 266:6626–6636, 1991
32. Passavant CW, Coulter SN, Wuepper KD: Structure and expression of the human keratolinin gene (abstr). Clin Res 38:225A, 1990
33. Rogers GE: Some aspects of the structure of the inner root sheath of hair follicles revealed by light and electron microscopy. Exp Cell Res 14:378–387, 1958
34. Rogers GE: Structural and biochemical features of the hair follicle. In: Montagna W, Lobitz W (eds.). The Epidermis. Academic Press Inc., New York, 1964, pp 179–236
35. Rogers GE: Isolation and properties of inner sheath cells of hair follicles. Exp Cell Res 33:264–276, 1964
36. Parakkal PF, Matoltsy AG: A study of the differentiation products of the hair follicle cells with the electron microscope. J Invest Dermatol 43:23–34, 1964
37. Steinert PM: Intermediate filaments (IF). In: Harris JR (ed.). Electron Microscopy of Proteins. Academic Press Inc., London, 1981, pp 123–166
38. Rothnagel JA, Rogers GE: Trichohyalin, an intermediate filament-associated protein of the hair follicle. J Cell Biol 102:1419–1429, 1986